

APPENDIX A

Statistical Analysis of Change in Exotic Species Presence in Association with the Rivers Edge 3D Seismic Survey

APPENDIX A

STATISTICAL ANALYSIS OF CHANGE IN EXOTIC SPECIES PRESENCE IN ASSOCIATION WITH THE RIVERS EDGE 3D SEISMIC SURVEY

Submitted to:

National Park Service
Big Thicket National Preserve
6044 FM 420
Kountze, TX 77625

Prepared For:

Cimarex Energy Company
1700 Lincoln Street; Suite 1800
Denver CO, 80203

and

Century Houston Exploration, Inc.
10210 Grogan's Mill Road
The Woodlands, TX 77380

Prepared By:

Alex Konomi and Betik Kyriakopoulos

Contracted to:

Dixie Environmental Services Co., LP (DESCO)
26902 Nichols Sawmill
Magnolia, TX 77355
281-252-9799

JULY 2013

STATISTICAL REPORT

Exotic species are invasive, non-native species to the region. Naturally, there is an introduction of exotic species as time passes. In addition, human intervention and severe natural phenomenon can accelerate the growth or assist in the introduction of exotic species. While there is no control over severe, natural phenomenon, we can try to control any adverse effects that human intervention might have. The Rivers Edge 3D seismic survey occurred between 2011 and 2013 in the Beaumont Unit and other lands administered by the Big Thicket National Preserve in southeast Texas. Due to the nature of the research (see **Section 2.4** for details) there are concerns that 3D seismic surveys could facilitate the increase presence of exotic species. The following report details the statistical analysis on this issue.

Problem:

Field data has been collected on the presence of exotic species (herein referred as dependent variable). In addition, other factors (herein referred as independent variables) that are expected to influence the presence of exotic species have been recorded as well. A total of 296 plots was re-examined in 2013 from the original 2011 survey of 391 plots. Additionally, qualitative notes were taken on the presence of a forest canopy gap (Gap) in all plots in 2011 and 2013. The question is if there is enough statistical evidence to suggest that any of the factors has influenced the presence of exotic species.

Segmentation Proposal:

The presence of exotic species was examined into two layers or strata, herbaceous (herb) and woody. Although the two layers originate from the same plot, they will be considered as independent experimental units in order to give more power to the statistical tests and hence error on the side of caution. The first part of the statistics report is separated into **Section 1: Herb Layer** and **Section 2: Woody Layer**. The presence of exotic species was also examined on a forest-wide scale, as viewed in the third segment of this report, **Section 3: Ranking Analysis**. All three sections of the statistics report examine the importance values for each exotic species as defined in **Section 2.2** of the Post-Operational Vegetation Assessment Report.

Section 1: Herb Layer

In an effort to establish the impact of the seismic research on the proliferation of exotic species, the statistical analysis for this layer needs to examine three key components: Time, Gap Presence, and Seismic Impact. The influence of each of the components will be measured with the Modified Importance Value (MIV) as defined in **Section 2.2**. As illustrated on figure 1, not only do each of these components appear influential, but also the interactions among them might be significant.

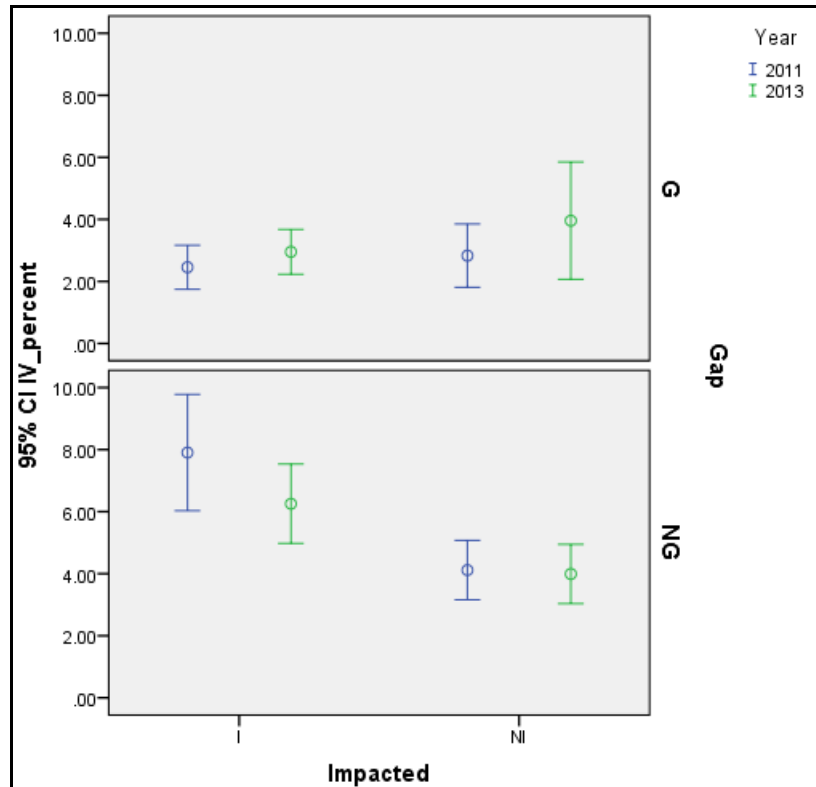


Figure 1: Mean and confidence interval for the impacted and non-impacted as a function of time for each of the Gap level (Gap and Non Gap).

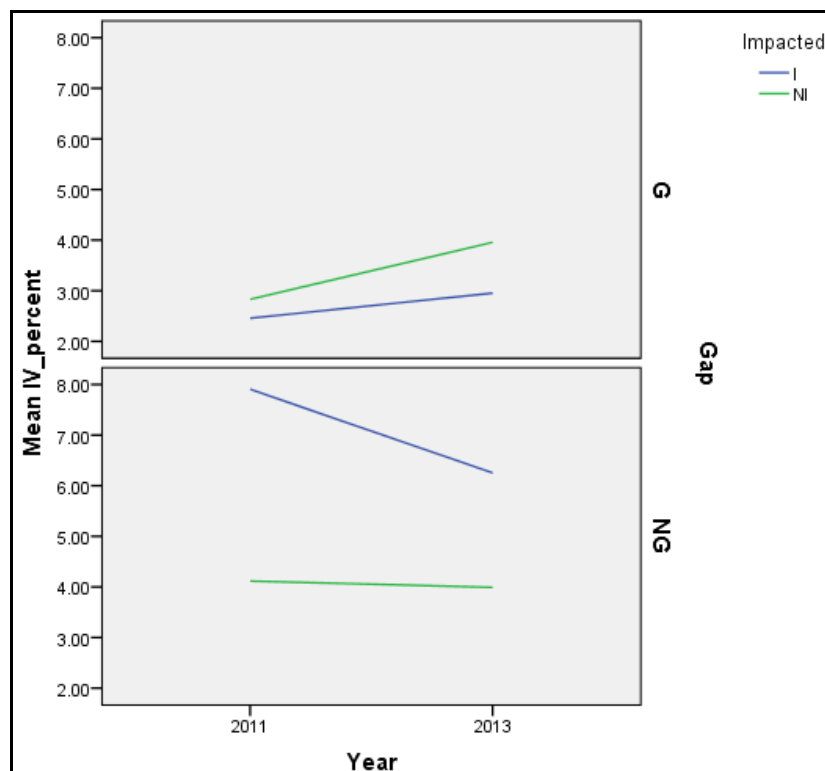


Figure 2: Mean of the impacted and non-impacted as a function of time for each of the Gap level.

Working hypothesis

Ho: From 2011 to 2013 there is no difference in the deviation of the presence of exotic species between areas impacted from seismic research (in 2012) vs. areas not impacted by seismic research (in 2012).

Ha: From 2011 to 2013 there is a significant difference in the deviation of the presence of exotic species in areas impacted from seismic research (in 2012) vs. areas not impacted from seismic research (in 2012).

Scientifically accepted type I error level $\alpha = 0.05$

Choosing the most powerful test:

Different testing methods were considered such as a t-test, ANOVA, and multiple linear regression (MLR) analysis on the MIV differences between 2011 and 2013. The repeated measure test with a diagonal compound symmetry covariance matrix of errors was chosen as the most powerful test.

Why use repeated measure test with a diagonal compound symmetry covariance matrix of errors?

Repeated Measures Analysis is used when there are variables of the same subject over time. In this scenario MIV is measured in the same plot over time (in 2011 and 2013). The usual analysis of ANOVA and the other previously mentioned statistical tools make the assumption of independence of the data. Because variables of same subject over time are dependent it is necessary to account for it. The next step is to decide what kind of dependence we have. The question in statistical terms translates to which covariance matrix of the error should be used? The answer to this question can vary based on the criteria that are used. The “best” criterion used widely in the statistic community is the AIC (Akaike information criterion). This criterion requires that the covariance matrix is specified. In this case “heterogeneous” indicates that the variances are different in the two time periods (2011 & 2013). “Compound symmetry” refers to the fact that the correlation remains the same over time.

In mathematical terms the final model will be the following

$$Y_{ijkl} = \mu + \tau_i + \beta_j + t_l + (\tau\beta)_{ij} + (\tau t)_{il} + (\beta t)_{jl} + d_{ijk} + e_{ijkl}$$

Y = indicates levels of MIV

μ = indicates a base level of exotic species. No matter the influence of humans or nature condition there will be some presence of exotic species.

τ_i = Main effect that accounts for the influence of the Seismic Impact

β_j = Main effect that accounts for the influence of Gap.

t_l = Time effect

$(\tau\beta)_{ij}$ = the two way interaction of Seismic Impact and Gap

$(\tau t)_{il}$ = the two way interaction of Time and seismic Impact

$(\beta t)_{jl}$ = the two interaction of Time and Gap

$(\tau\beta t)_{ijl}$ = the three way interaction of Seismic Impact, Gap and Time

$(d)_{ijk}$ = random effect of each plot (subject) which is an observational unit

ε_{ijk} = error term accounting for randomness on each individual observation

We estimate the parameters of the model described above using restricted maximum likelihood (REML) approach. The REML approach gives more stable estimations and is preferred when we have repeated measure data. The likelihood function is calculated from a transformed set of data so that the nuisance parameters have no effect. Moreover, since we do not have missing cells we use the Type III method. The Type III sums of squares have one major advantage in that they are invariant with respect to the cell frequencies as long as the general form of the estimability remains constant among all the cells. This type of sums of squares is often considered the best for an unbalanced model with no missing cells which is the case in our application.

Following the above we compute the ANOVA table.

Table 1

Type III Test of Fixed Effects ^a				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	351.428	252.444	.000
Year	1	351.182	.030	.862
Impacted	1	351.428	4.574	.033
Gap	1	372.069	21.087	.000
Year * Impacted	1	351.182	5.470	.020
Year * Gap	1	351.661	13.059	.000
Impacted * Gap	1	372.069	12.719	.000
Year * Impacted * Gap	1	351.661	1.102	.295

Based on the results from **Table 1** we conclude the following:

Intercept IS statistically significant (p-value=0.0001< alpha=0.05). This means that no matter if other factors are present or not, there will always be some kind of exotic species presence.

3-way interaction of time-gap-impact is NOT statistically significant (p-value=0.295 > alpha=0.05). This means that interaction of Year and Seismic Impact is not significant in the presence of Gap. Equivocally, Time and Gap are not significant in the presence of Seismic Impact and Gap/ Seismic Impact combination is not significant in the presence of Time.

2-Way interaction of Seismic Impact and Gap IS statistically significant (p-value=0.0001< alpha=0.05). This means that depending on the presence of canopy gap or not the seismic Impact will have a different influence on the presence of exotic species.

2-Way interaction of Time and Gap IS statistically significant (p-value=0.0001< alpha=0.05). This means that depending on the presence of canopy gap or not the time will have a different influence on the presence of exotic species.

2-Way interaction of Seismic Impact and Time IS statistically significant (p-value=0.020<alpha=0.05). This means that depending on if plots were impacted or not by the Seismic Impact Time will have a different influence on the presence of exotic species.

Main effects of Time, Gap and Seismic does not carry a statistical meaning given that their corresponding interactions are significant.

Overall conclusion for the Herbal layer is that when there is a Gap in the canopy of a plot the Seismic Impact does not have an influence on the presence of exotic species. However, when there is no canopy gap present in the plot, the impacted plots experience a statistically significant reduction in presence of exotic species. It must be noted that the presence of exotic species in the plots that were going to get impacted was higher in 2011 than the control plots in 2011. Therefore, it can be argued that the reduction in presence of exotic species in 2013 for the impacted plot can be a simple regression to the mean, rather than an effect of the Seismic Impact. To test this hypothesis a new experiment must be completed, but it is beyond the scope of this analysis as a reduction of exotic species is not an adverse effect.

Supplementary material

To better understand the above analysis we also give the 95% confidence intervals statistics of the IV percentage mean for each of the levels differences we present the Impacted, Gap and Year.

Table 2

Year	Impacted	Gap	Mean	Std. Error	df	95% Confidence Interval	
						Lower Bound	Upper Bound
2011	I	G	2.457	.541	472.779	1.394	3.520
		NG	7.905	.638	472.779	6.651	9.159
	NI	G	2.913	.615	501.528	1.705	4.120
		NG	4.053	.544	495.039	2.984	5.122
2013	I	G	2.953	.541	472.779	1.890	4.016
		NG	6.254	.638	472.779	5.000	7.508
	NI	G	4.000	.622	502.123	2.777	5.223
		NG	3.960	.539	494.453	2.901	5.019

Section 2: Woody Layer

In an effort to establish the impact of the Seismic Research, the statistical analysis for this layer needs to examine three key components: Time, Gap Presence, Seismic Impact. The influence of each component will be measured with the Relative Importance Value (RIV) as defined on **Section 2.2**. As illustrated on **Figure 1**, there seems to be significant components and potential interactions.

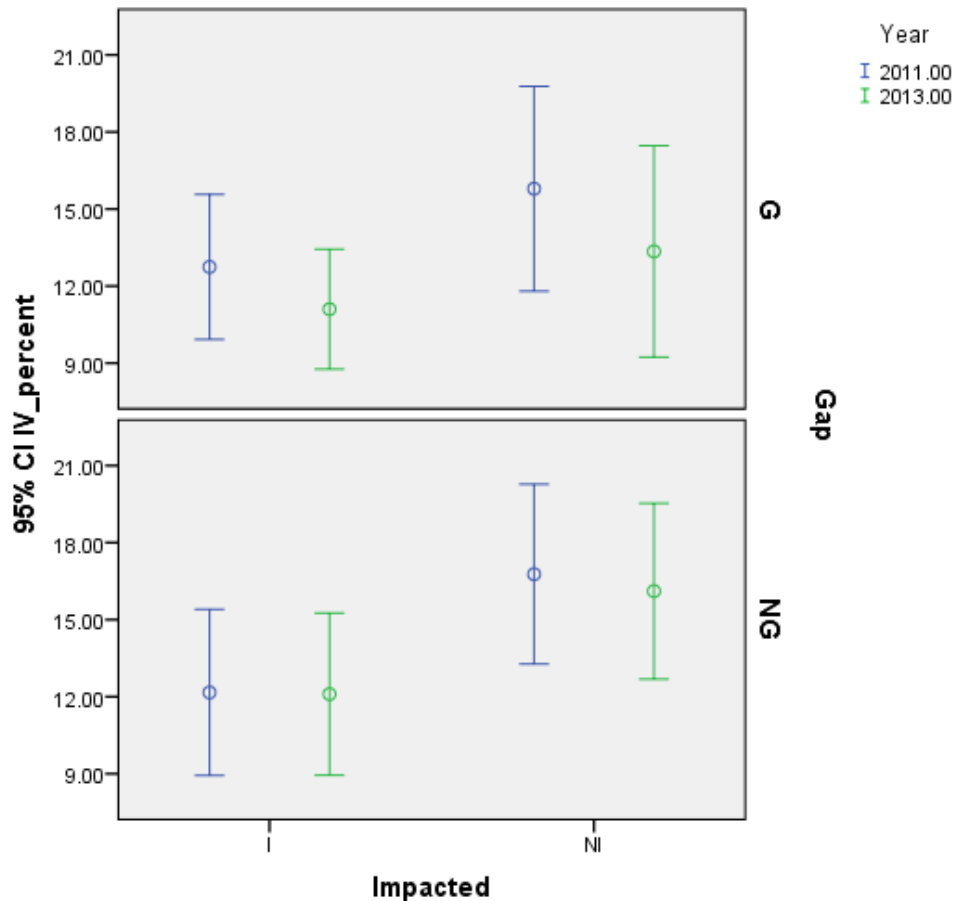


Figure 3: Mean and confidence interval for the impacted and non-impacted as a function of time for each of the Gap level (Gap and Non Gap).

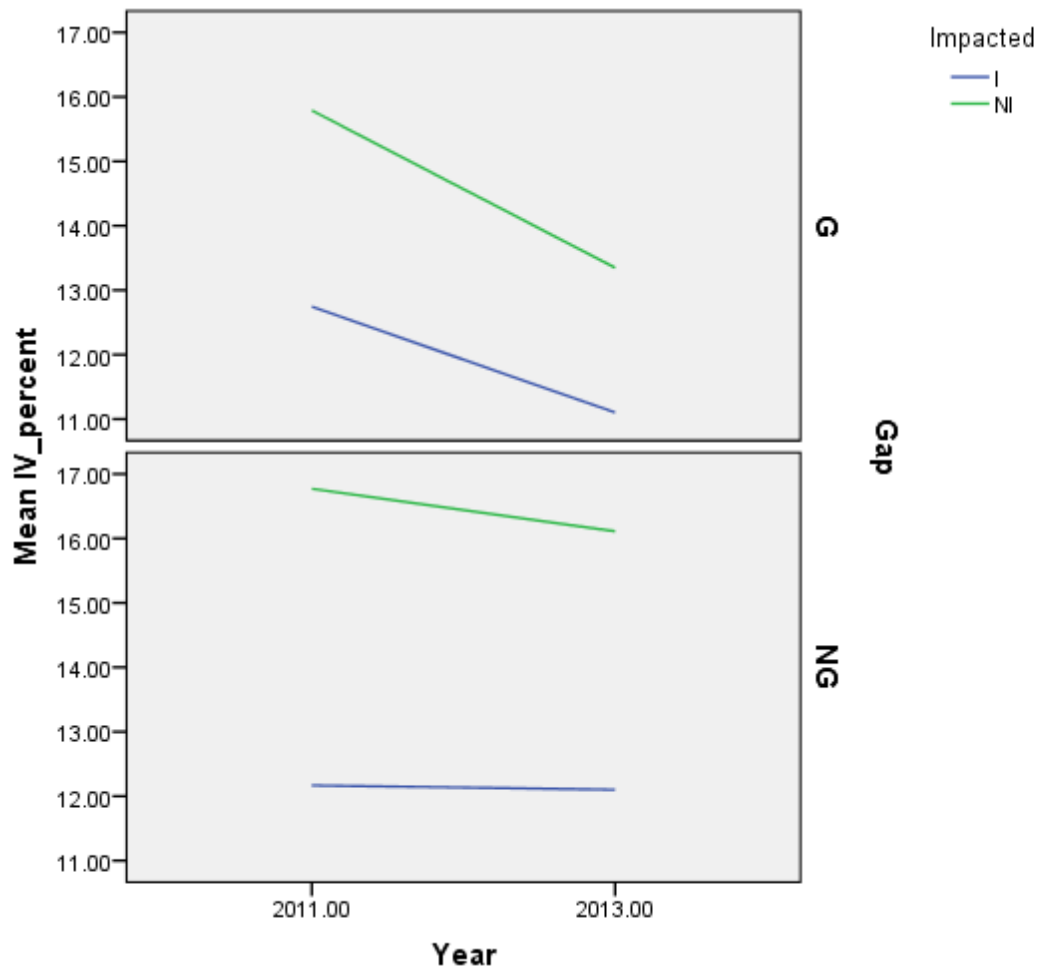


Figure 4: Mean of the impacted and non-impacted as a function of time for each of the Gap level.

Working hypothesis

Ho: From 2011 to 2013 there is no difference in the deviation of the presence of exotic species between areas impacted from seismic research (in 2012) vs. areas not impacted by seismic research (in 2012).

Ha: From 2011 to 2013 there is a significant difference in the deviation of the presence of exotic species in areas impacted from seismic research (in 2012) vs. areas not impacted from seismic research (in 2012).

Scientifically accepted type I error level $\alpha = 0.05$

Choosing the most powerful test:

Different testing methods were considered such as a t-test, ANOVA, and multiple linear regression (MLR) analysis on the RIV differences between 2011 and 2013. The repeated measure test with a diagonal compound symmetry covariance matrix of errors should be chosen as the most powerful test.

Why use repeated measure test with a diagonal compound symmetry covariance matrix of errors?

See Herbal layer explanation as the design and sampling techniques are statistically identical.

As before, in mathematical terms the final model will be the following.

$$Y_{ijk} = \mu + \tau_i + \beta_j + t_l + (\tau\beta)_{ij} + (\tau t)_{il} + (\beta t)_{jl} + d_{ijk} + e_{ijk}$$

Y_{ijk} = indicates RIV levels of for each individual k at different Time, Seismic Impact level and Gap level.

μ = indicates a base level of exotic species. No matter the influence of humans or nature condition there will be some presence of exotic species.

τ_i = Main effect that accounts for the influence of the Seismic Impact

β_j =Main effect that accounts for the influence of Gap

t_l = Time effect

$(\tau\beta)_{ij}$ = the two way interaction of Seismic Impact and Gap

$(\tau t)_{il}$ = the two way interaction of Time and Seismic Impact

$(\beta t)_{jl}$ = the two interaction of Time and Gap

$(\tau\beta t)_{ijl}$ = the three way interaction of Seismic Impact , Gap and Time

$(d)_{ijk}$ = random effect of each plot (subject) which is an observational unit

e_{ijk} = error term accounting for randomness on each individual observation

Following the same estimation strategies as in the Herbal analysis we compute the ANOVA table with the F-test statistics and their corresponding p-values.

Table 3

ANOVA Table				
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	194.311	290.759	.000
Impacted	1	194.311	4.631	.033
Gap	1	211.648	.452	.502
Year	1	194.155	5.707	.018
Impacted * Gap	1	211.648	.290	.591
Impacted * Year	1	194.155	.472	.493
Gap * Year	1	194.175	2.508	.115
Impacted * Gap * Year	1	194.175	.000	.987

a. Dependent Variable: I.V. [%].

Based on the results from **Table 3** we conclude the following:

Intercept IS statistically significant (p-value=0.0001 < $\alpha=0.05$). This means that no matter if other factors are present or not, there will always be some kind of exotic species presence.

3-way interaction of time-gap-impact is NOT statistically significant (p-value=0.295 > $\alpha=0.05$). This means that interaction of Year and Seismic Impact is not significant in the presence of Gap. Equivocally, Time and Gap are not significant in the presence of Seismic Impact and Gap Seismic Impact combination is not significant in the presence of Time.

2-Way interaction of Seismic Impact and Gap IS NOT statistically significant (p-value=0.591 > $\alpha=0.05$). This means that Seismic Impact does not need be interpreted in terms of canopy gap presence or not.

2-Way interaction of Time and Gap IS NOT statistically significant (p-value=0.115 > $\alpha=0.05$). This means that Time does not need be interpreted in terms of canopy gap presence or not.

2-Way interaction of Seismic Impact and Time IS NOT statistically significant (p-value=0.493 > $\alpha=0.05$). This means that Seismic Impact does not need be interpreted in terms of Time.

Main effect of Canopy Gap presence IS NOT statistically significant (p-value=0.502 > $\alpha=0.05$). This means that Canopy Gap has no influence on the presence of exotic species.

Main effect of Time IS statistically significant (p-value=0.018 < $\alpha=0.05$). This means that as time elapses the presence of exotic species changes.

Main effect of Seismic Research Impact IS statistically significant (p-value=0.033 < $\alpha=0.05$). This means that the RIV mean of the plots where Seismic Impact has been applied is different from the RIV mean of the plots where no-Seismic Impact has been applied.

All the interaction terms are statistically insignificant and can be removed from the model with an F-test. Moreover, the main effect of Gap level is also statistically insignificant (p-value=0.502 > 0.1 = importance level).

The overall conclusion for the Woody layer is that from 2011 to 2013 there is a statistically significant decrease on presence of exotic species. This is also evident from the control (not impacted plots). The seismic impact influence is also statistically significant in the sense that impacted plots have lower presence of exotic species than non-impacted plots. This is not the same as saying that the Seismic Impact reduced the presence of exotic species. As indicated on figure 4 impacted plots (in 2011 & 2013) have significant lower presence of exotic species.

Supplementary material

To better explain the above analysis we also give the 95% confidence intervals statistics of the IV percentage mean for each of the levels differences we present the Impacted, Gap and Year.

Table 4							
Impacted	Gap	Year	Mean	Std. Error	df	95% Confidence Interval	
						Lower Bound	Upper Bound
I	G	2011.00	12.745	1.480	231.323	9.829	15.660
		2013.00	11.101	1.480	231.323	8.186	14.017
	NG	2011.00	12.167	1.942	231.323	8.341	15.994
		2013.00	12.101	1.942	231.323	8.275	15.928
NI	G	2011.00	15.720	1.682	254.141	12.407	19.032
		2013.00	13.371	1.700	254.596	10.023	16.719
	NG	2011.00	16.831	1.557	250.547	13.763	19.898
		2013.00	16.091	1.544	250.103	13.051	19.132

Section 3: Ranking Analysis

In addition to RIV/MIV value on a given plot, species are also given an overall forest-wide RIV/MIV value. Based on this overall RIV/MIV value each species takes a ranking value among all species (native and not native). This rank can then be used to determine the forest-wide hierarchical position of a given species and therefore the position of an exotic species.

Woody Layer

The ranking analysis for this layer will examine if the ranks in 2013 follow the same trend as the rank in 2011. The underlying assumption here is that if a species maintained the same rank (or relatively the same rank) in the two time periods then it is reasonable to assume that the population distribution has remained relatively constant.

Ho: The Ranking in 2013 is independent of the Ranking in 2011 (i.e. 2013 ranking does not follow 2011 ranking)

Ha: The Ranking in 2011 is predictable of the Ranking in 2013 (i.e. 2013 ranking follows 2011 ranking)

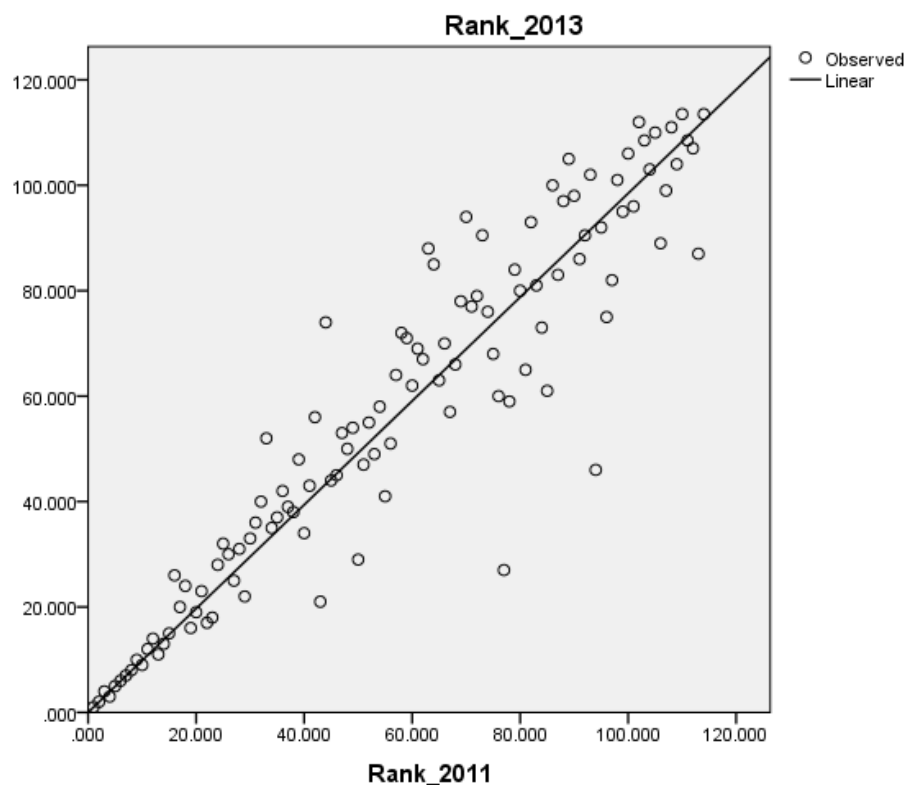


Figure 3: Ranks between 2011 and 2013 indicate a rather low variability with minor changes among the ranks.

A linear regression methodology is used in order to determine if 2011 ranking is predicable of 2013 ranking.

Table 1

ANOVA ^{a,b}					
Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	485082.509	1	485082.509	3587.092	.000
Residual	15280.991	113	135.230		
Total	500363.500 ^d	114			

a. Dependent Variable: Rank_2013

b. Predictors: Rank_2011

Table 2

Model Summary ^a			
R	R Square	Adjusted R Square	Std. Error of the Estimate
.985	.969	.969	11.629

a. Independent Variable is Rank_2011.

Table 3

Table 3								
Coefficients ^{a,b}								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	Rank_2011	.985	.016	.985	59.892	.000	.952	1.017

a. Dependent Variable: Rank_2013

b. Linear Regression through the Origin

Table 4

Woody Layer	2013		2011	
	RIV (%)	Rank	RIV (%)	Rank
<i>Triadica sebifera</i> (triseb)	10.15	1	11.44	1
<i>Lonicera japonica</i> (lonjap)	0.05	96	0.04	103
<i>Ligustrum japonicum</i> (ligjap)	0.04	99.5	N/A	N/A

Based on the results from **Tables 1, 2, 3, and 4** we conclude the following:

The 2011 Ranking IS statistically significant (p-value=0.0001< alpha=0.05). This means that on average the 2011 rank can predict the 2013 rank. The slope is estimated at 0.985 which indicates a very close association between the 2011 and 2013 rankings. In other words the slope says that on average if a species ranked 5 in 2011, it will rank 4.925 in 2013. The 95% confidence interval of this “average” species is (4.76, 5.36). Note that this is a prediction on the average among all species; not a prediction for an individual species. Furthermore, the order of the ranking for the exotic species has remained the same between 2011 & 2013 (see table 4).

The overall conclusion for the Woody layer is that the Ranks in 2013 are very close to the Ranks from 2011. In general terms the variability among the ranks is rather low (correlation =95.8%). It must be noted that the variability has the tendency increase as the rank values increase (i.e. smaller footprint - meaning low RIV values-). This is consistent with expectations as trees with rather large trunk are less likely to change in a 2 year period than trees with a rather small trunk.

Herbal Layer

The ranking analysis for this layer will examine if the ranks in 2013 follow the same trend as the rank in 2011. The underlying assumption here is that if species maintained the same rank (or relatively the same rank) in the two time periods then it is reasonable to assume that the population distribution has remained relatively constant.

Ho: The Ranking in 2013 is independent of the Ranking in 2011 (i.e. 2013 ranking does not follow 2011 ranking)

Ha: The Ranking in 2011 is predictable of the Ranking in 2013 (i.e. 2013 ranking follows 2011 ranking)

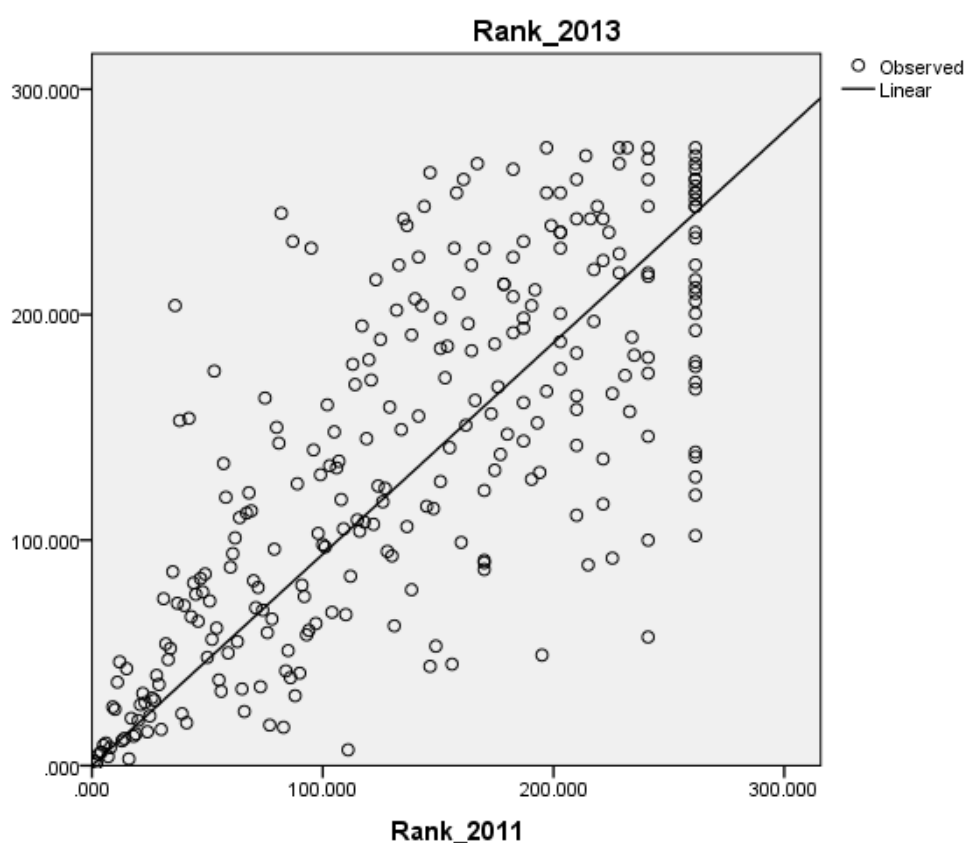


Figure 2: Ranks between 2011 and 2013 indicate a rather high variability with drastic changes among the high ranks.

A linear regression methodology is used in order to determine if 2011 ranking is predicable of 2013 ranking.

Table 5

ANOVA ^a					
	Sum of Squares	Df	Mean Square	F	Sig.
Regression	6193396.453	1	6193396.453	1997.019	.000
Residual	852863.047	275	3101.320		
Total	7046259.500	276			

a. The independent variable is Rank_2011.

Table 6

Coefficients ^{a,b}							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
1 Rank_2011	.938	.021	.938	44.688	.000	.896	.979

a. Dependent Variable: Rank_2013

b. Linear Regression through the Origin

Table 7

Model Summary ^a			
R	R Square	Adjusted R Square	Std. Error of the Estimate
.938	.879	.879	55.689

The independent variable is Rank_2011.

a. The equation was estimated without the constant term.

Based on the results from **Tables 5, 6, and 7** we conclude the following:

The 2011 Ranking IS statistically significant (p-value=0.0001< alpha=0.05). This means that on average the 2011 rank can predict the 2013 rank. The slope is estimated at 0.938 which indicates a very close association between the 2011 and 2013 rankings. In other words the slope says that on average if a species ranked 5 in 2011, it will rank 4.69 in 2013. The 95% confidence interval of this “average” species is (4.48, 4.89). Note that this is a prediction on the average among all species; not a prediction for an individual species.

Although the results appear similar to the woody layer, we should exercise caution on jumping to conclusions as the variability is rather high. Tables 8 and 9 list the top 20 species and their ranks exchanges between 2011 vs. 2013 and vice versa (2013 vs. 2011). Table 10 lists the ranks for the exotic species.

Table 8: Top Twenty Ranked Species in 2011 and the Corresponding Rank in 2013

2013			2011		
Species Name	Percentage	Rank	Species Name	Percentage	Rank
<i>Alternanthera philoxeroides</i>	3.40%	2	<i>Alternanthera philoxeroides</i>	8.14%	1
<i>Salvinia minima</i>	5.45%	1	<i>Salvinia minima</i>	7.57%	2
<i>Zizaniopsis miliacea</i>	2.25%	5	<i>Zizaniopsis miliacea</i>	5.57%	3

Table 8: Top Twenty Ranked Species in 2011 and the Corresponding Rank in 2013

2013			2011		
Species Name	Percentage	Rank	Species Name	Percentage	Rank
<i>Chasmanthium laxum</i>	2.16%	6	<i>Chasmanthium laxum</i>	2.92%	4
<i>Vitis rotundifolia</i>	1.82%	9	<i>Vitis rotundifolia</i>	2.82%	5
<i>Ilex vomitoria</i>	1.81%	10	<i>Ilex vomitoria</i>	2.59%	6
<i>Toxicodendron radicans</i>	2.87%	4	<i>Toxicodendron radicans</i>	1.98%	7
<i>Rubus argutus</i>	1.88%	8	<i>Rubus argutus</i>	1.87%	8
<i>Bignonia capreolata</i>	1.05%	26	<i>Bignonia capreolata</i>	1.40%	9
<i>Ptilimnium capillaceum</i>	1.05%	25	<i>Ptilimnium capillaceum</i>	1.33%	10
<i>Callicarpa americana</i>	0.68%	37	<i>Callicarpa americana</i>	1.31%	11
<i>Phanopyrum gymnocarpon</i>	0.56%	46	<i>Phanopyrum gymnocarpon</i>	1.30%	12
<i>Triadica sebifera</i>	1.66%	11	<i>Triadica sebifera</i>	1.29%	13
<i>Campsis radicans</i>	1.60%	12	<i>Campsis radicans</i>	1.20%	14
<i>Rhynchospora corniculata</i>	0.61%	43	<i>Rhynchospora corniculata</i>	1.12%	15
<i>Liquidambar styraciflua</i>	2.98%	3	<i>Liquidambar styraciflua</i>	1.11%	16
<i>Ilex opaca</i>	1.09%	21	<i>Ilex opaca</i>	1.04%	18
<i>Carpinus caroliniana</i>	1.48%	13	<i>Carpinus caroliniana</i>	1.04%	19
<i>Parthenocissus quinquefolia</i>	1.44%	14	<i>Parthenocissus quinquefolia</i>	1.02%	20
<i>Brunnichia ovata</i>	1.10%	20	<i>Brunnichia ovata</i>	1.01%	21

Table 9: Top Twenty Ranked Species in 2011 and the Corresponding Rank in 2013

2013			2011		
Species Name	Percentage	Rank	Species Name	Percentage	Rank
<i>Salvinia minima</i>	5.45%	1	<i>Salvinia minima</i>	7.57%	2
<i>Alternanthera philoxeroides</i>	3.40%	2	<i>Alternanthera philoxeroides</i>	8.14%	1
<i>Liquidambar styraciflua</i>	2.98%	3	<i>Liquidambar styraciflua</i>	1.11%	16
<i>Toxicodendron radicans</i>	2.87%	4	<i>Toxicodendron radicans</i>	1.98%	7
<i>Zizaniopsis miliacea</i>	2.25%	5	<i>Zizaniopsis miliacea</i>	5.57%	3
<i>Chasmanthium laxum</i>	2.16%	6	<i>Chasmanthium laxum</i>	2.92%	4
<i>Galium tinctorium</i>	2.08%	7	<i>Galium tinctorium</i>	0.16%	120
<i>Rubus argutus</i>	1.88%	8	<i>Rubus argutus</i>	1.87%	8
<i>Vitis rotundifolia</i>	1.82%	9	<i>Vitis rotundifolia</i>	2.82%	5
<i>Ilex vomitoria</i>	1.81%	10	<i>Ilex vomitoria</i>	2.59%	6
<i>Triadica sebifera</i>	1.66%	11	<i>Triadica sebifera</i>	1.29%	13
<i>Campsis radicans</i>	1.60%	12	<i>Campsis radicans</i>	1.20%	14
<i>Carpinus caroliniana</i>	1.48%	13	<i>Carpinus caroliniana</i>	1.04%	19
<i>Parthenocissus quinquefolia</i>	1.44%	14	<i>Parthenocissus quinquefolia</i>	1.02%	20
<i>Dichanthelium commutatum</i>	1.42%	15	<i>Dichanthelium commutatum</i>	0.88%	25
<i>Oplismenus hirtellus</i>	1.40%	16	<i>Oplismenus hirtellus</i>	0.71%	31
<i>Symphyotrichum dumosum</i>	1.23%	17	<i>Symphyotrichum dumosum</i>	0.24%	89
<i>Ilex decidua</i>	1.22%	18	<i>Ilex coriacea</i>	0.25%	82
<i>Tradescantia ohioensis</i>	1.14%	19	<i>Tradescantia ohioensis</i>	0.50%	42
<i>Brunnichia ovata</i>	1.10%	20	<i>Brunnichia ovata</i>	1.01%	21

Table 9: Top Twenty Ranked Species in 2011 and the Corresponding Rank in 2013

2013			2011		
Species Name	Percentage	Rank	Species Name	Percentage	Rank
<i>Ilex opaca</i>	1.09%	21	<i>Ilex opaca</i>	1.04%	18

Table 10: Rank of Exotic Species

Herbal Layer	2013		2011	
	MIV (%)	Rank	MIV (%)	Rank
<i>Alternanthera philoxeroides</i> (altphi)	3.4	2	8.14	1
<i>Triadica sebifera</i> (triseb)	1.29	13	1.66	11
<i>Lonicera japonica</i> (lonjap)	0.98	30	0.85	27
<i>Lygodium japonicum</i> (lygjap)	0.38	64	0.40	56
<i>Salvinia molesta</i> (salmol)	0.39	63	N/A	N/A
<i>Eichhornia crassipes</i> (eiccra)	0.25	95	0.04	248
<i>Ligustrum japonicum</i> (ligjap)	0.11	151	N/A	N/A
<i>Cyperus entrerianus</i> (cypent)	0.01	440.5	N/A	N/A

In an effort to better understand the forest wide presence of species in the herbal layer, we will also analyse the MIV values in addition to the rank. This analysis is deemed necessary given the large variability observed on the ranks between 2011 & 2013 (see figure 2).

Percentage presence Analysis (herbal layer)

The statistical methodology here is based on the linear regression; however, we transform the percentage of the species such that they can take values between $(-\infty, +\infty)$. This can be achieved with a logit transformation ($\text{logit}(p) = \log(p/(1-p))$).

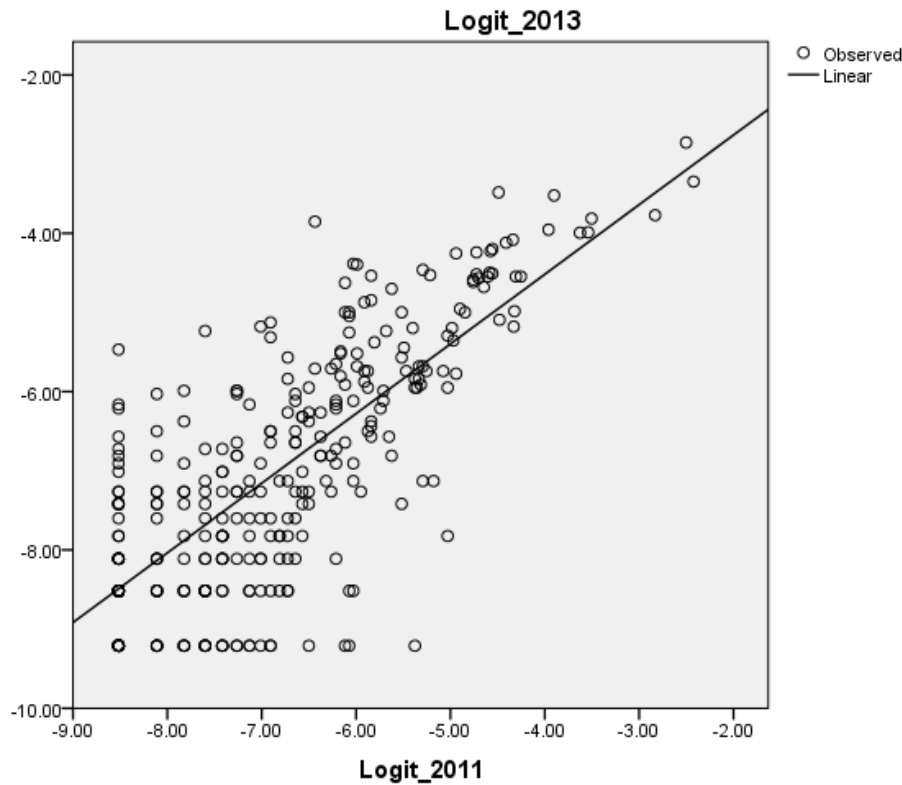


Figure 3: Logit points between 2011 and 2013

Ho: The percent MIV values in 2013 are not correlated to the percent MIV values in 2011 (i.e. 2013 MIV values do not follow 2011 MIV values)

Ha: The percent MIV values in 2013 are correlated to the percent MIV values in 2011 (i.e. 2013 MIV values follow 2011 MIV values)

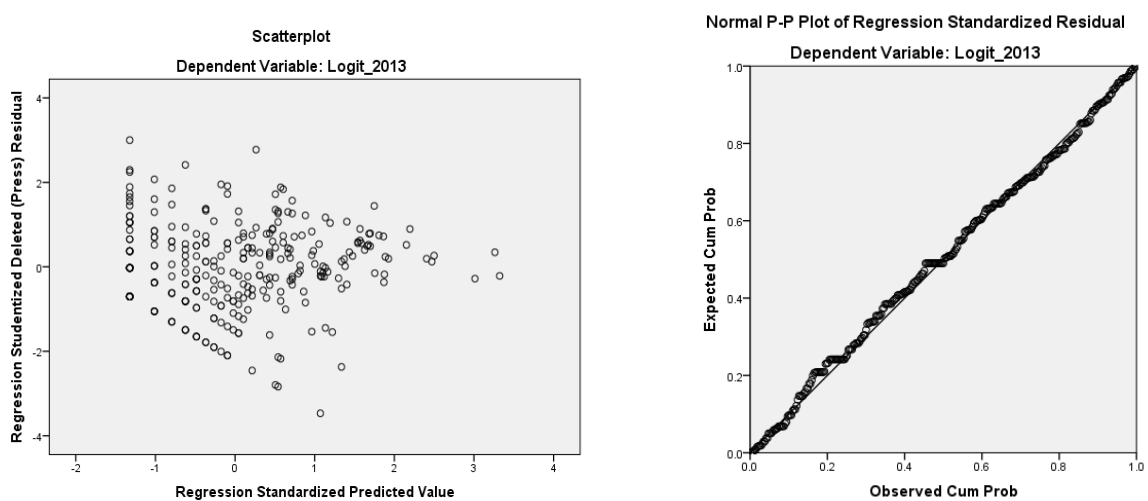


Figure 4: Residual plots show the diagnostic assumptions required to run the regression into the logit transformation – a categorical data analysis of the percentage.

Table 11

ANOVA ^a						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	404.601	1	404.601	384.843	.000 ^b
	Residual	318.557	303	1.051		
	Total	723.158	304			

a. Dependent Variable: Logit_2013

b. Predictors: (Constant), Logit_2011

Table 12

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta			Lower Bound	Upper Bound
(Constant)	-1.000	.310		-3.229	.001	-1.609	-.391
Logit_2011	.880	.045	.748	19.617	.000	.791	.968

a. Dependent Variable: Logit_2013

Table 13

Model Summary ^b				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.748 ^a	.559	.558	1.02535

a. Predictors: (Constant), Logit_2011

b. Dependent Variable: Logit_2013

Based on the results from **Tables 11, 12, and 13** we conclude the following:

The 2011 percent MIV IS statistically significant (p-value=0.0001< alpha=0.05). This means that on average the 2011 MIV values are correlated to 2013 MIV values (Correlation coefficient R= 74.8%). Given that logit based estimates are meaningful only on a log odds space, their estimation/interpretation has been omitted as it goes beyond the scope of this project.

Based on the results from **Table 10** we conclude that the order of ranking among exotic species has remained the same for 2011 & 2013. There are two noteworthy observations from table 10. 1

1. Significant decrease of “*Alternanthera philoxeroides*” MIV values (from 8.14% in 2011 to 3.40% in 2013). This decrease needs to be addressed in context with other exotic species which has been addressed on the impact analysis. In terms of rank “*Alternanthera philoxeroides*” moved from Rank 1 to 2 in 2013 which is well within the margin for randomness.
2. Three new exotic species that were not observed in 2011 are now visible in 2013. Although, this is an adverse event, we must not focus on the species while missing the forest. In other words, all those 3 species have ranks above 50 and rather low MIV values (below 0.50%). This “discovery” in 2013 falls well within random chance. Hence, it is

safe to assume that rankings of exotic species have remained rather stable from 2011 to 2013.

The overall conclusion for the Herbal layer is that the Ranks in 2013 follow the trend observed in 2011 and that the MIV values in 2013 are correlated to MIV values in 2011. Nevertheless, it is hard to argue that 2011 can explain 2013 among all species and the entire rank spectrum. Table 13 shows that forest wide only 55.8% of the variability in 2013 can be explained based on 2011. However, on table 10 we observe that when it comes to exotic species, 2011 is a rather good indicator of 2013 as the ranks between 2011 and 2013 have remained virtually the same. Although, it can be said with certainty that there is a positive correlation between 2011 and 2013, the determination of “same” or “different” can only be based on an operational point scenario.